

# Master Data Cleansing Services

First step towards effective Master Data Management



*Clean, consistent and accurate Master Data provides the foundation for maximizing the ROI and for completing meaningful spend analysis. Prospecta Software assists in creating and maintaining clean material master, customer master and vendor master files.*

Prospecta Software specializes in providing data cleansing. It employs a panel of domain experts who work with operational teams in order to arrive at right noun-modifier combination, its attributes and the correctness of attribute values. Prospecta Software can provide attribute extraction to any industry standards or it can develop its own set of attributes using internal domain knowledge.

***Prospecta Software offers complete range of data cleansing services, which include:***

#### **Noun Modifier Identification**

Noun modifiers help in unambiguously identifying a product and helps in easy grouping the products. Prospecta Software can assign Noun Modifiers to products with the help of existing descriptions.

#### **Data Normalization**

We format the data entities such as Manufacturer/supplier names, attributes values, abbreviations, Unit of measurements etc to a uniform and consistent values based on a specific Style-guide.

#### **De-Duplication**

Duplicates are identified and removed by adopting various permutations of pattern matching rules, database queries and manual inspection.

#### **Attribute Value extraction**

We provide attribute value extraction to any industry standard attribute schema and also develop new set of attribute by using in-house domain expert's knowledge.

## Identify and qualify the master data and its sources

The definition of master data may be different for different business units. The first step involves identifying and qualifying master data for each business unit in the organization. For example, from an accounts receivables perspective, the customer address is the master data, whereas from a distribution perspective, location is the master data. It is then necessary to identify the source systems/applications which store/generate the master information. Perform detailed analysis of the source system's master data structure, and map between the source systems and MasterDataOnline.

## Identify and define the global and local data elements

More than one system may store/generate the same master information. Additionally, there could also be a global version as well as local versions of the master data. Perform detailed analysis to understand the commonalities and differences between local, global and global-local attributes of data elements. This analysis is critical for an organization that has distributed global and regional applications and deals with multilingual and multicurrency scenarios. Based on this analysis, a right data sourcing strategy for a MasterDataOnline is defined.

## Identify the data elements that require data cleansing and correction

At this stage, the data elements supporting the MasterDataOnline that require data cleansing and correction have to be identified. For example, a party record would comprise - name, address ZIP code, contact phone number, demographic profile, psychographic profile, etc. It is important to scope these data elements for data quality. Communication with the stakeholders is necessary so that as part of the MDM initiative, data quality will be injected into these selected data elements on an organization-wide basis.

## Perform data discovery and analysis

Data collected from source applications needs to be analyzed to understand the sufficiency, accuracy, consistency and redundancy issues associated with data sets. Analyze source data from both business and technical perspectives. The data profiling and data quality analysis reports reflect the complete diagnosis of the health of data sets like:

- Basic data statistics and frequency analysis (patterns, unique count, occurrences, etc.).
- Missing and duplicate attributes of the master data (name and address analysis, etc.).
- Incorrect and out-of-range value analysis.
- Data profiling and analysis as per predefined business and technical rules.
- Cross comparison of data elements between source systems.
- Data irregularity analysis (heterogeneous spelling, mixed case, etc.).

It is strongly advised to use a data profiling tool to analyze datasets. Data analysis should focus on technical as well as business metadata, and should be quantitative to report data quality.

# Introduction

## PROSPECTA

Prospecta Software is an SAP Services Partner and was established in the year 2002 with focus on building innovative complementary solutions to meet the business gaps across industry verticals. Prospecta works closely with its partners to ensure that its clients receive the best of services & solutions and build strong partnerships to offer complete solutions. Prospecta believes in valuing the partners as valuing the customers. Prospecta shares a very strong relationship with SAP across Asia Pacific Region.

---



### Prospecta Soft Pvt. Ltd.

Prospecta House, K-2/5, DLF City, Phase II, Gurgaon – 122002.  
Phone: +91-124 - 4754500 / 4754534

### Prospecta Software Australia Pty. Ltd.

Suite 507/5, Celebration Drive, Norwest Business Park, Baulkham Hills NSW 2153.  
Phone: +61-2-8824-9749

For more information contact us at [sales@prospectasoftware.com](mailto:sales@prospectasoftware.com) or visit us @ [www.prospectasoftware.com](http://www.prospectasoftware.com)